GA-ReLU: an activation function for Geometric Algebra Networks applied to 2D Navier-Stokes PDEs

Alberto Pepe[1], Sven Buchholz[2], Joan Lasenby[1]

[1] Signal Processing & Communications Lab, University of Cambridge

[2] Department of Computer Science and Media, Technical University Brandenburg

Introduction

- Networks in **Geometric Algebra (GA)** are very effective at capturing the geometry of data.
- Data are embedded as **multivectors** of a given algebra.
- Inputs, weights and biases to the network are all multivectors.
- **Q: how to apply activation functions to multivectors?**

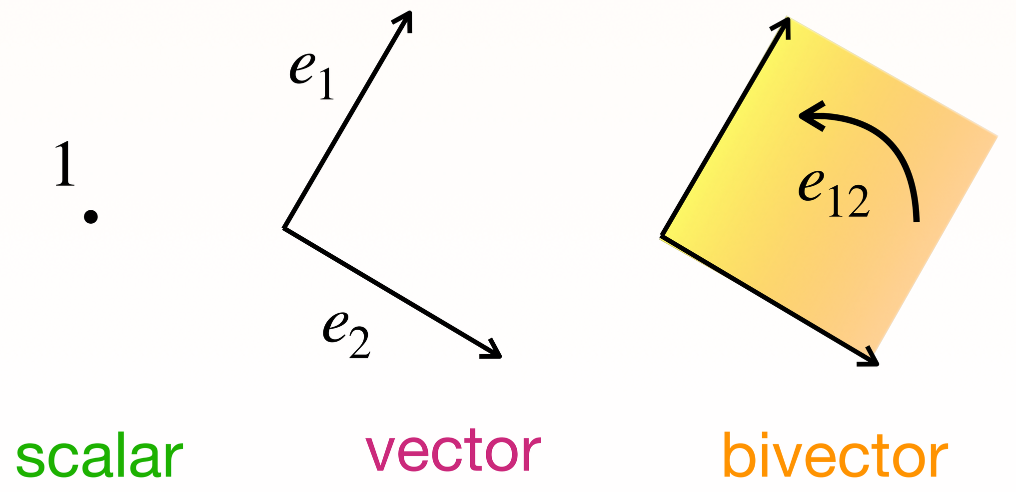

Fig.1: elements in 2D GA.

Problem

- It is hard to define a function theory for multivectors.
- Generally, a real-valued activation function ϕ is applied over each real coefficient of the multivector bases independently.

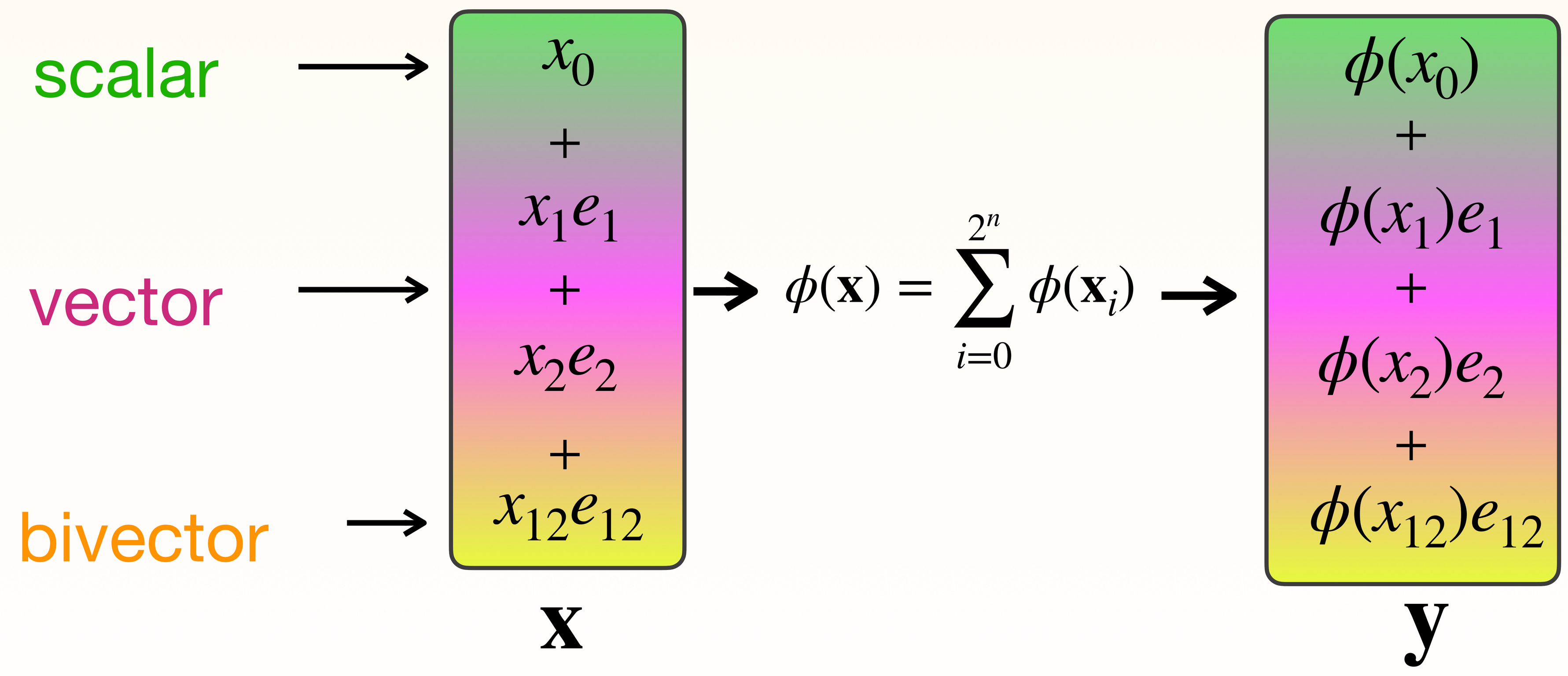

Fig.2: element-wise activation function over multivector **x**.

- This approach does not take into account interactions between elements, and **y** does not have the same geometrical meaning of **x**.
- **We propose GA-ReLU to keep the activation function grounded in geometry.**

Methodology

- We look at the 2D Navier-Stokes problem over a square domain.
- We embed smoke (scalar) and velocity (vector) as multivectors in G(2,0,0).
- **GA-ReLU** ψ is the composition of ϕ, element-wise ReLU, and *f*, complex ReLU.

$$\psi(\mathbf{x}) = (\phi \circ f)(\mathbf{x}) = \phi(x_0) + \phi\left(K\left(\tan^{-1}\left(\frac{x_2}{x_1}\right)\right)x_1\right)e_1 + \phi\left(K\left(\tan^{-1}\left(\frac{x_2}{x_1}\right)\right)x_2\right)e_2$$

- ϕ keeps the magnitude bounded;
- *f* introduces a dependency on the vector phase.

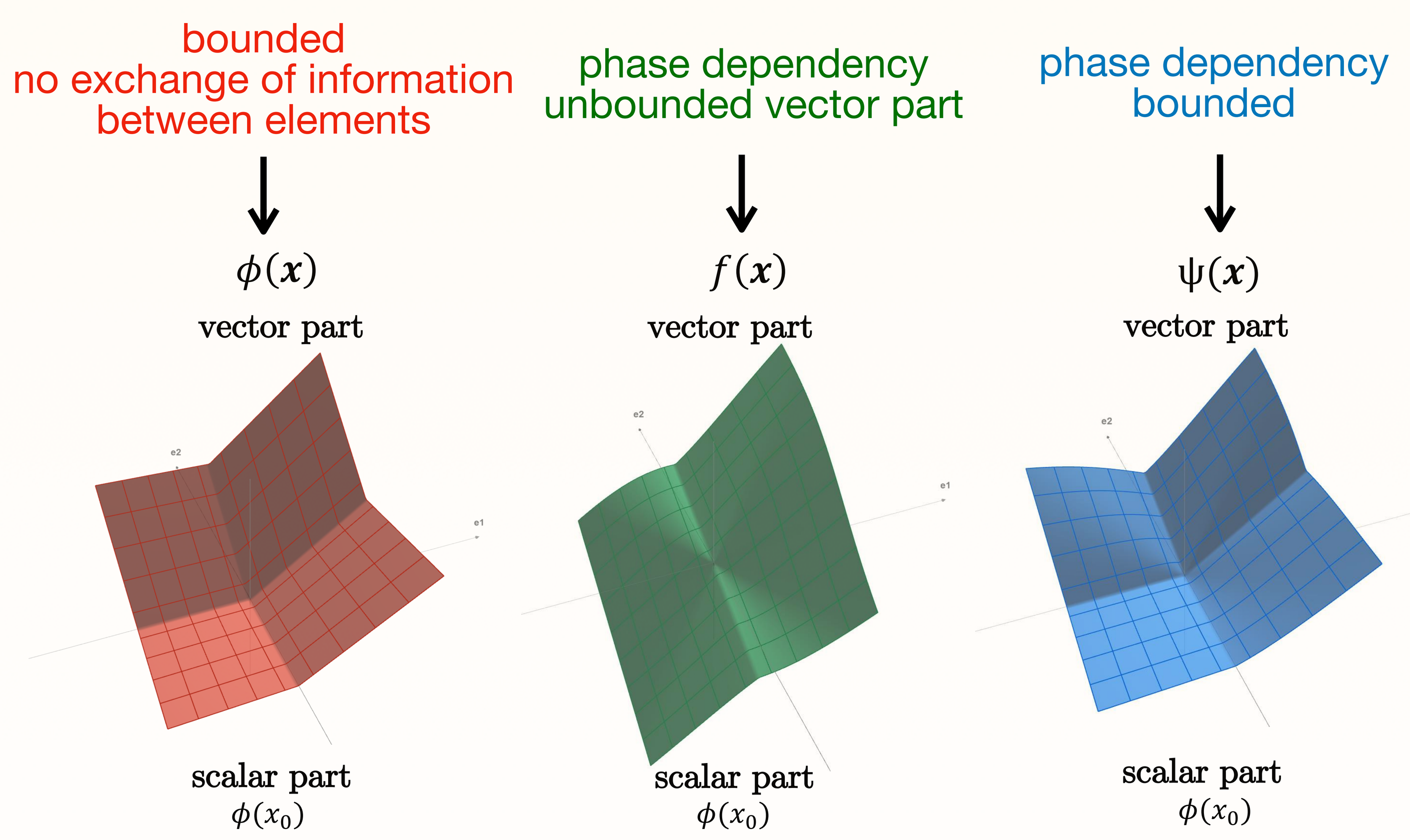

Fig.4: element-wise ReLU, complex ReLU, GA-ReLU.

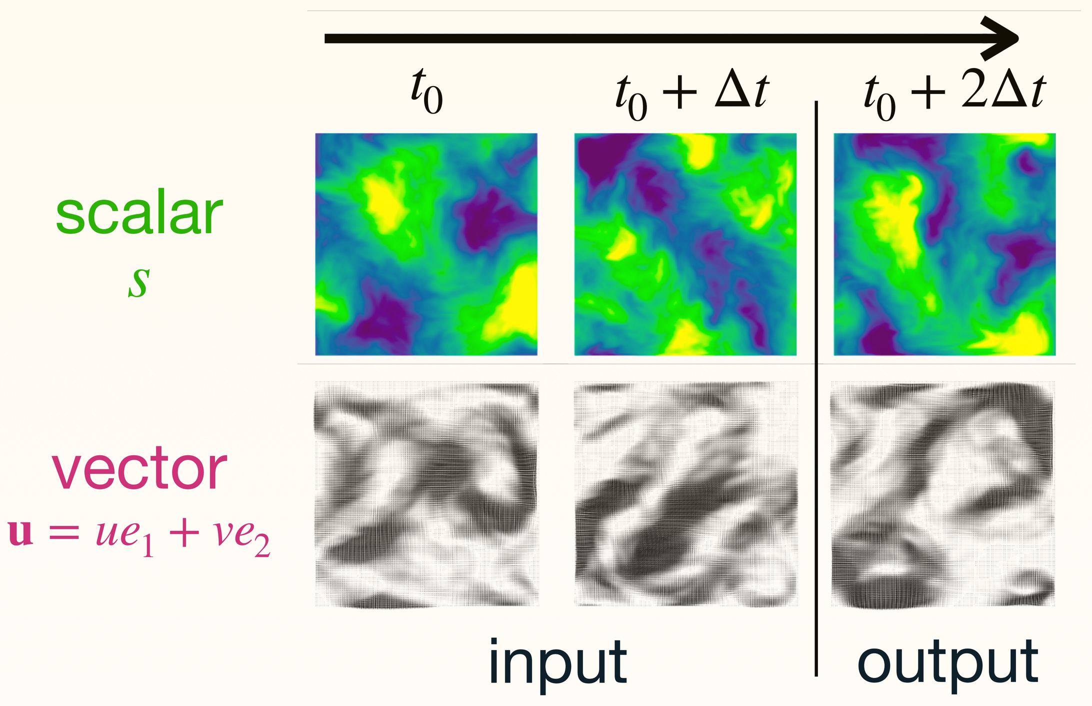

Fig.3: embedding PDE variables as multivectors.

- The complex ReLU is defined as:

$$f(z) = \frac{1}{2}\left(1 + \cos(\arg(z))\right) z = K(\arg(z))z$$

- We rewrite our multi vector in 2D as a sum of two complex numbers:

$$\mathbf{x} = (x_0 + Ix_{12}) + e_1(x_1 + Ix_2) = z_S + e_1 z_V$$

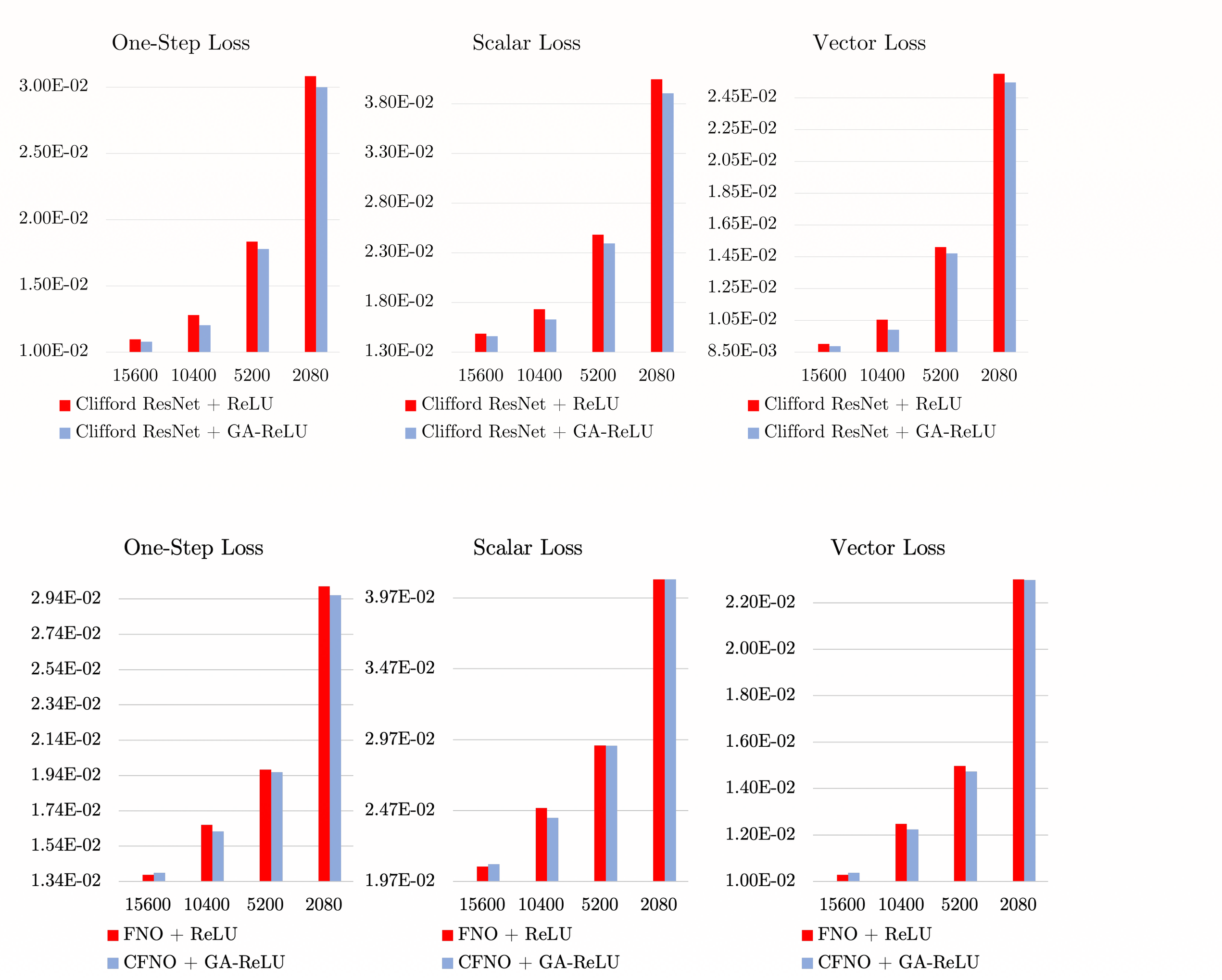

Fig.5: Errors versus number of training data for Clifford ResNet and Clifford FNO with ReLU and GA-ReLU activation functions.

Results

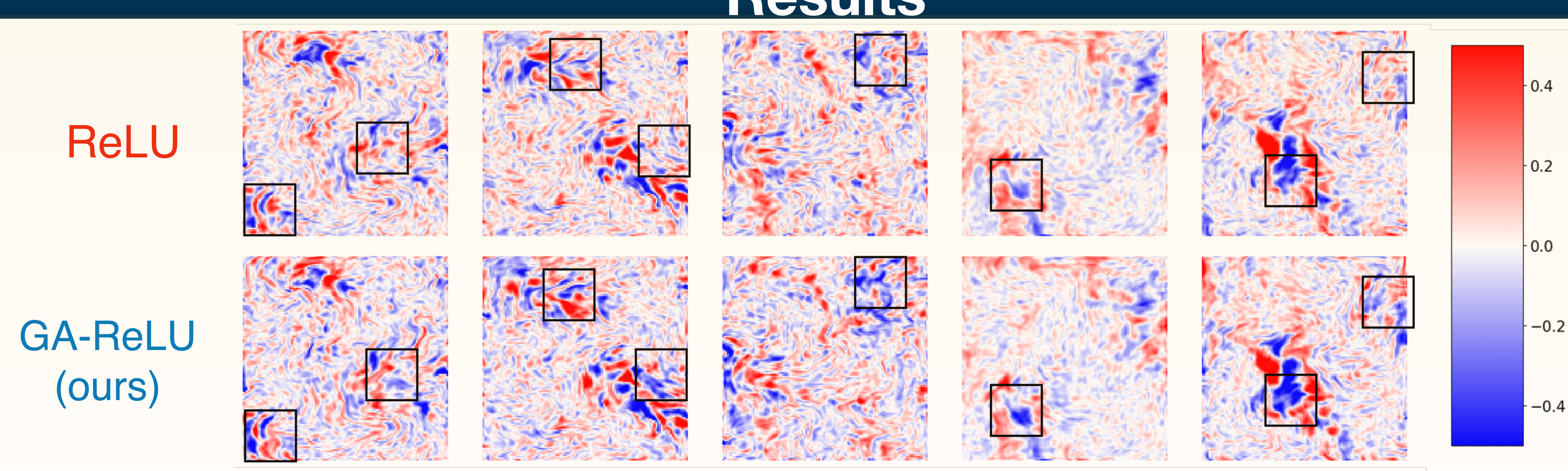

Fig. 6: Difference between ground truth and predicted scalar fields for 5 different time instants.

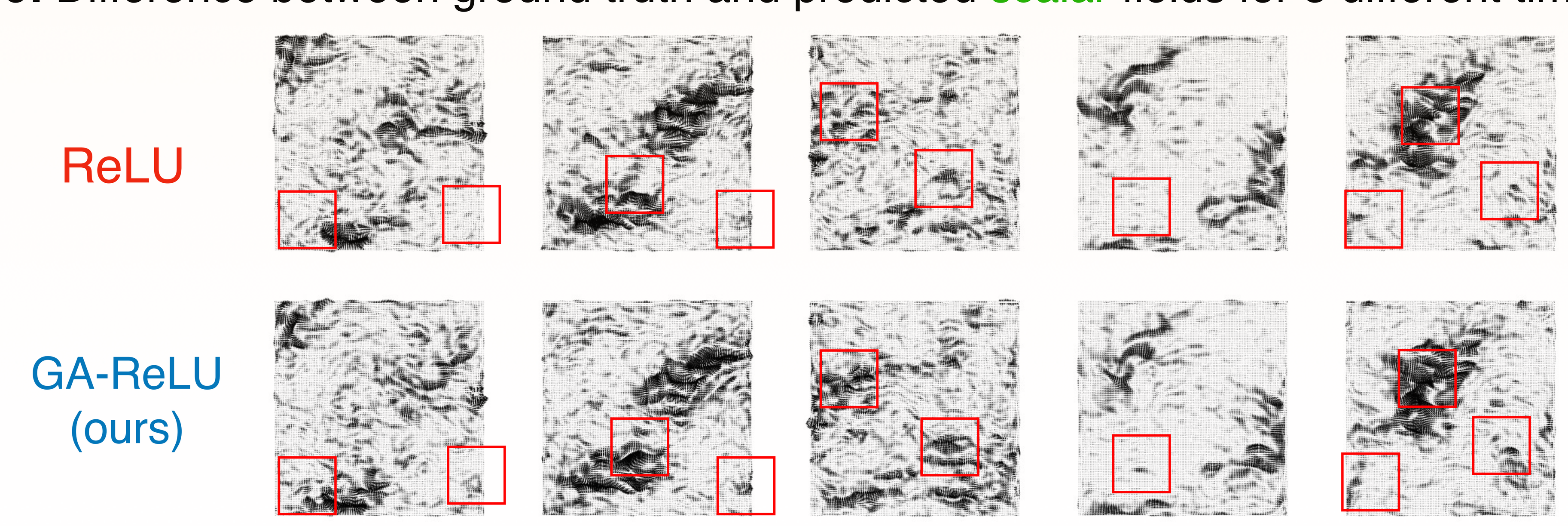

Fig. 7: Difference between ground truth and predicted vector fields for 5 different time instants.

Contributions

- Design of a geometry-informed activation function for G(2,0,0)
- More accurate PDE solutions with minimal intervention on the activation function.

ICLR

ICLR 2024 Workshop on AI4DifferentialEquations in Science

ap2219@cam.ac.uk